THE ROYAL SOCIETY

How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies

J. A. J. Metz[1,2*] and M. Gyllenberg[3]

[1]*Institute for Evolutionary and Ecological Sciences, Leiden University, NL-2311 GP Leiden, The Netherlands*
[2]*Adaptive Dynamics Network, IIASA, A-2361 Laxenburg, Austria*
[3]*Department of Mathematics, University of Turku, Fin-20014 Turku, Finland*

We define a fitness concept applicable to structured metapopulations consisting of infinitely many equally coupled patches. In addition, we introduce a more easily calculated quantity R_m that relates to fitness in the same manner as R_0 relates to fitness in ordinary population dynamics: the R_m of a mutant is only defined when the resident population dynamics converges to a point equilibrium and R_m is larger (smaller) than 1 if and only if mutant fitness is positive (negative). R_m corresponds to the average number of newborn dispersers resulting from the (on average less than one) local colony founded by a newborn disperser. Efficient algorithms for calculating its numerical value are provided. As an example of the usefulness of these concepts we calculate the evolutionarily stable conditional dispersal strategy for individuals that can account for the local population density in their dispersal decisions. Below a threshold density $\tilde{x}$, at which staying and leaving are equality profitable, everybody should stay and above $\tilde{x}$ everybody should leave, where profitability is measured as the mean number of dispersers produced through lines of descent consisting of only non-dispersers.

Keywords: evolutionary ecology; fitness; dispersal; metapopulation

1. INTRODUCTION

During the past decade there has been an explosion of interest in the dynamics of metapopulations. The book edited by Hanski & Gilpin (1997) and the one by Hanski (1999) each contain more than 1000 references. At the same time the evolution of dispersal has caught the interest of many scientists. As most species have a hierarchical spatial structure with many local populations connected by dispersal comprising a metapopulation, the evolution of dispersal is most conveniently studied in the technical framework of metapopulation dynamics (Hastings 1983; Holt & McPeek 1996; Doebeli & Ruxton 1997; Jánosi & Scheuring 1997; Parvinen 1999; Travis *et al.* 1999). An excellent discussion of the literature on life-history calculations in metapopulations, including evolutionarily stable (ES) dispersal strategies, can be found in Olivieri & Gouyon (1998; see also Travis *et al.* 1999; Ronce *et al.* 2000).

Most metapopulation models assume an infinite number of equally coupled patches. This poses some problems when it comes to performing evolutionarily stable strategy (ESS) or adaptive dynamics (Dieckmann & Law 1996; Metz *et al.* 1996; Geritz *et al.* 1997, 1998) calculations for these models. It is not immediately clear how we should define fitness and what would be useful stand-ins for fitness that are comparable to R_0 in the single population case. In this paper we solve these problems. In order to show the usefulness of the concepts introduced, we calculate the ES conditional dispersal strategy if individuals can sense the local population density, thus solving a problem posed by Olivieri & Gouyon (1998; see also Ezoe & Iwasa 1997; Jánosi & Scheuring 1997; Travis *et al.* 1999).

Before embarking on our programme we provide a little context. In reality we have a finite though possibly large number of patches with different characteristics which are coupled in a complicated manner. One of the simplifying assumptions often made in metapopulation theory, to which we shall also adhere, is that all patches have equal characteristics, though not equal population sizes and are equally coupled (all patches exchange individuals on an equal footing). The assumptions of equal patches can easily be removed (Hanski & Gyllenberg 1993; Gyllenberg *et al.* 1997). We adhere to it in order to avoid notational clutter. The assumption of equal coupling is essential. However, for many purposes it provides a fair approximation to reality. As an example consider aphids with the patches corresponding to single leaves on a tree. Our educated guess is that there will rarely be any harm in applying our results in cases where the root-mean-square dispersal distance is more than three times the interpatch distance: for sufficiently wide dispersal the correlation between local densities will be low (compare the various contributions in Dieckmann *et al.* (2000)) and the stream of arrivals is determined by an average over the dispersal from a largish number of patches, resulting in a rapid convergence to a local mean field model. Moreover, results for mean field models about invasion into point equilibria immediately extend to homogeneous equilibria in local mean field models.

Given the homogeneity assumption just introduced, we have two overall system parameters available, namely size of the patches ω and number of patches Ω. When both

Proc. R. Soc. Lond. B (2001) **268**, 499–508
Received 22 May 2000   *Accepted* 26 September 2000

499

parameters are small the metapopulation becomes extinct on an ecological time-scale due to demographic stochasticity. Therefore, it only makes sense to study long-term evolution if at least one of these parameters is large. We shall take the mathematician's stance and equate large with infinite so that we may rigorously neglect process properties that at large sizes effectively disappear from sight.

In case the patch sizes (ω) are infinite and there are few patches (Ω small) the classical fitness concept for structured populations, as expounded by Metz *et al.* (1992), applies. If there are no further structuring variables we have Ω local population densities as the state variables of the metapopulation. The same applies to a rare invader. Thus, the local linearization of the invader dynamics near zero invader densities yields an Ω-dimensional linear dynamics, possibly with time-varying coefficients, depending on whether the overall environment is constant and the resident population dynamics converge to a point attractor or some more complicated resident population dynamics ensues. The dominant Lyapunov exponent of this linear dynamics (biologically, the long-term time averaged per-capita growth rate), in the case of constant coefficients the dominant eigenvalue, provides the right fitness concept for being inserted into evolutionary calculations (e.g. Parvinen 1999). In the case of a population dynamical point equilibrium, we can use the general R_0 concept for structured populations introduced by Diekmann *et al.* (1990, 1998; see also Heesterbeek 1992) as a stand-in for fitness: if and only if $R_0 > 1$ (<1) will fitness be positive (negative). Since the sign of its fitness determines whether a mutant can invade, knowledge of this sign suffices for many evolutionary arguments including the calculation of ESSs. Therefore, in those cases it is sufficient to calculate R_0, which is usually easier, provided the conditions that are necessary for its definition are satisfied.

The challenge comes when Ω is infinite. Then we are outside the standard framework for establishing the existence of an exponential growth rate which, according to the arguments in Metz *et al.* (1992), is how we should define fitness. Nor can we follow some standard approach for calculating an R_0-like quantity.

In order to keep the mathematics simple we concentrate on the case without further structuring variables. In addition, we assume that only newborns disperse and that they do so by entering a dispersal pool which they leave either by dying or by entering a patch. We phrase the models such that the newborns are allowed to disperse with a probability dependent on the local population density and, on encountering a patch, choose to stay or to leave again depending on the population density that they encounter. (Compare with Ruxton & Rohani (1999): not only is juvenile dispersal the natural strategy for sessile organisms, it also prevails among mobile ones, probably since there is a cost involved in settling that juveniles have to pay regardless.) The reason for explicitly taking account of such conditional decision rules is that this directly leads to our closing example. However, the arguments below are of a general nature and not tied to those specific assumptions. A companion paper (Gyllenberg & Metz 2001) dissects the abstract, mathematical structure of the argument within the general framework for structured population models put forward

in Diekmann *et al.* (1994, 1998, 2001) and calculates the ES conditional dispersal strategy for adult dispersal.

2. THE FINITE PATCH SIZE CASE

(a) *Model specification*

The state equations for a structured metapopulation are similar to the differential equations for the state probabilities of a continuous-time Markov process. There are good reasons for this similarity. Every single patch undergoes a Markov process. Only the collective of all infinitely many patches together behaves as a deterministic entity. In the case of finite ω the Markovian state of a single patch (p state) corresponds to the number of individuals present in it. To make life easy we assume that there is a maximum k to the number of individuals in a patch. The state of the metapopulation (m-state) is given by the fractions p_i of patches with various numbers (i) of individuals in them together with the disperser density D. Figure 1 shows the possible p-state transitions for the model we have in mind. α is the rate constant of patch encounter for the dispersers, μ_i are per-capita death rates, s_i are probabilities that a newly arrived immigrant decides to stay, λ_i are per-capita birth rates, d_i are probabilities that a newly born individual decides to disperse and γ_i are rates of externally imposed catastrophes wiping out the local population. The strategy parameters d_i and s_i are assumed to be heritable. The other parameters (α, γ_i, μ_i, and λ_i) are supposed to be either constants or fluctuate in an ergodic manner. (In practice this means that α, γ_i, μ_i and λ_i should not show any systematic trends on any relevant time-scale. The reason for invoking the term 'ergodic' is to guarantee the truth of some of our mathematical statements below.) The m-state equations corresponding to the scheme in figure 1 are as follows.

$$\frac{\mathrm{d}p_0}{\mathrm{d}t} = -\alpha D s_0 \, p_0 + \mu_1 \, p_1 + \sum_{j=1}^{k} \gamma_j \, p_j, \tag{1a}$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = [\alpha D s_{i-1} + (i-1)\lambda_{i-1}(1-d_{i-1})]p_{i-1}$$
$$- [i\mu_i + \alpha D s_i + i\lambda_i(1-d_i) + \gamma_i]p_i + (i+)\mu_{i+1} \, p_{i+1}, \tag{1b}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = -\alpha \sum_{i=0}^{k-1} p_i s_i D + \sum_{i=1}^{k} i\lambda_i d_i \, p_i - \mu_D D, \tag{1c}$$

where μ_D is the per-capita death rate of the dispersers, which is also assumed to be either constant or ergodic. We used the conventions that $p_{-1} = 0$, $p_{k+1} = 0$, $s_k = 0$ and $d_k = 1$ (the first convention only becomes relevant at a later stage) in equation (1). The first two conventions bring the form of the equations for the p_i with i next to the boundary of the feasible domain in line with those for the p_i with i in the interior of that domain. The assumptions about s_k and d_k are required to keep the population size from becoming larger than k. (Otherwise, if k is the absolute maximum to the local population size, anybody born or arriving into a population of size k had better leave lest they die.)

No general results about the equilibria of equation (1) are known. We know of no cases yet where it has been found that, for constant parameters, equation (1) have no
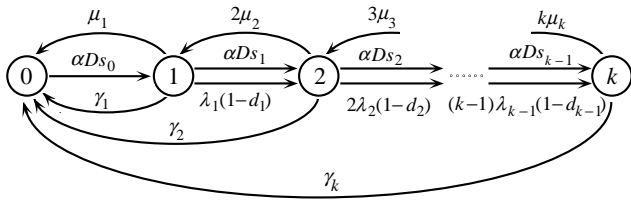
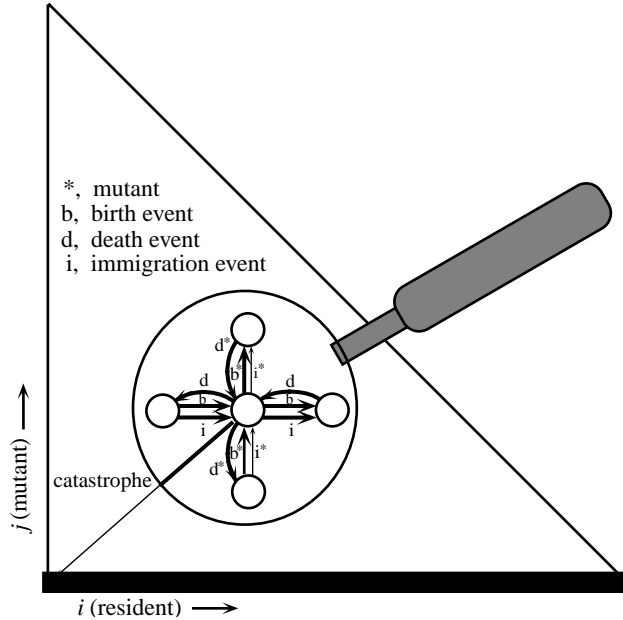Figure 1. State transition diagram for a single patch.

Figure 2. Schematic representation of the p-state transitions in a combined resident plus mutant patch. Thin arrows indicate transitions that are relatively rare.

globally stable point equilibrium. However, we surmise that it is possible to obtain bistability by incorporating a sufficiently strong Allee-type effect. Appendix A provides an algorithm for calculating the equilibrium values of the relative frequencies of the local population sizes $\widehat{p}_i$ and of the disperser density $\widehat{D}$.

(b) *The linearized mutant equations*

Now consider what happens when a mutant having strategy parameters d^* and s^* is introduced. Then we have to consider an extended set of state variables for the metapopulation, for which we choose the relative frequencies of the patches filled with different resident and mutant numbers $q_{i,j}$, $i \geqslant 0$, $j \geqslant 0$ and $i+j \leqslant k$, where i refers to the number of resident-type individuals and j refers to the mutant number. In addition, we have equations for the resident and mutant disperser pools. The former we shall again denote as D and the latter as D^*. A schematic representation of the p-state transitions is indicated in figure 2.

We are primarily interested in the mutant population. This population is represented by those $q_{i,j}$ for which $j \neq 0$. Initially these $q_{i,j}$ may be supposed to be very small relative to the $q_{i,0}$. When the mutant population is still rare it has little influence on the resident dynamics. Therefore, $q_{i,0}$, $i = 0, \ldots, k$, can be approximated using p_i calculated from equation (1) after the resident

dynamics have relaxed to an attractor. When the environment is constant we substitute $q_{i,0} = \widehat{p}_i$ and $D = \widehat{D}$.

Initially, D^* will also be small. This allows us to neglect any further mutants arriving in the rare patches that already have mutants in them. We start with introducing some conventions in order to simplify the notation. We shall notionally set $q_{-1,j} = 0$ and $q_{i,j} = 0$ when $i+j = k+1$ and $s_{i,j} = s_{i,j}^* = 0$ and $d_{i,j} = d_{i,j}^* = 1$ whenever $i+j = k$. The rationale is the same as that for the similar conventions for p_i. Moreover, we shall add asterisks to all parameters pertaining to mutant individuals. For the concrete example under consideration, where mutants and residents differ only with respect to the probabilities that newborns disperse from a patch d_i and the probabilities that dispersers stay in a patch s_i write

$$\alpha^* = \alpha, \quad \lambda_{i,j}^* = \lambda_{i,j} = \lambda_{i+j}, \quad \mu_{i,j}^* = \mu_{i,j} = \mu_{i+j}, \quad \gamma_{i,j}^* = \gamma_{i+j},$$
$$d_{i,j}^* = d_{i+j}^*, \quad s_{i,j}^* = s_{i+j}^*, \quad d_{i,j} = d_{i+j}, \quad s_{i,j} = s_{i+j}.$$

The differential equations for $q_{i,1}$ have a term corresponding to the arrival of mutants into patches with residents only. In the equations for $q_{i,j}$, $j > 1$, this term is lacking.

$$\frac{dq_{i,1}}{dt} = \alpha^* D^* s_i^* p_i + [\alpha D s_{i,0} + (i-1)\lambda_{i,0}(1 - d_{i,0})] q_{i-1,1}$$
$$- [i\mu_{i,1} + \mu_{i,1}^* + \alpha D s_{i,1} + i\lambda_{i,1}(1 - d_{i,1}) + \lambda_{i,1}^*(1 - d_{i,1}^*)$$
$$+ \gamma_{i,1}^*] q_{i,1} + (i+1)\mu_{i+1,1} q_{i+1,1} + 2\mu_{i,2}^* q_{i,2}, \qquad (2a)$$

$$\frac{dq_{i,j}}{dt} = [\alpha D s_{i-1,j} + (i-1)\lambda_{i-1,j}(1 - d_{i-1,+j})]q_{i-1,j} + (j-1)$$
$$\times \lambda_{i,j-1}^*(1 - d_{i,j-1}^*)q_{i,j-1} - [i\mu_{i,j} + j\mu_{i,j}^* + \alpha D s_{i,j}$$
$$+ i\lambda_{i,j}(1 - d_{i,j}) + j\lambda_{i,j}^*(1 - d_{i,j}^*) + \gamma_{i,j}^*]q_{i,j} + (i+1)$$
$$\times \mu_{i+1,j}q_{i+1,j} + (j+1)\mu_{i,j+1}^*q_{i,j+1}, \qquad (2b)$$

$$\frac{dD^*}{dt} = -\alpha^* \sum_{i=0}^{k-1} p_i s_{i,0}^* D^* + \sum_{i=0}^{k-1}\sum_{j=1}^{k-i} j\lambda_{i,j}^* d_{i,j}^* q_{i,j} - \mu_D^* D^*. \quad (2c)$$

Equation (2) forms a sufficient starting point for the discussion of invasion fitness in structured metapopulations, even in the most general case where all the individual parameters are functions of some strategy parameter S, for example $\lambda_{i,j} = \Lambda_{i,j}(S, S^*)$, $\lambda_{i,j}^* = \Lambda_{j,i}(S^*, S)$ and $\lambda_i = \lambda_{i,0}$, with $\Lambda_{i,j}(S, S)$ not dependent on i and j separately but only on $i+j$ and $\Lambda_{i,0}(S, S^*)$ not dependent on S^*.

(c) *Invasion fitness*

Prior to the definition of fitness we rewrite equation (2) in a more accessible form. We thus define the (column) vector $\boldsymbol{V}$ as

$$\boldsymbol{V} = [(q_{i,j})_{i=0, \ldots, k-1, j=1, \ldots, k-i}, D^*]^{\mathrm{T}},$$

where the pairs (i,j) are supposed to be lexicographically ordered, i.e. put in the order $(0, 1)$, $(0, 2)$, ..., $(0, k)$, $(1, 1)$, $(1, 2)$, ..., $(1, k-1)$, ..., $(k-1, 1)$. The map transforming a pair (i,j), $i = 0, \ldots, k-1$, $j = 1, \ldots, k-i$, into its position n in the lexicografic order we shall call L. Using this notation we write

$$\frac{d\boldsymbol{V}}{dt} = B\boldsymbol{V}, \qquad (3)$$

where the matrix B contains the various coefficients from equation (2) ordered appropriately. If the coefficients in equation (2) fluctuate ergodically, there exists a dominant Lyapunov exponent ρ associated with equation (3). In biological terms, the total mutant population size overall will grow or decline exponentially at a per-capita rate ρ. Generalizing from the case of ordinary population dynamics we shall call ρ the fitness of the strategy $S^* = (d^*, s^*)$ in the environment set by an $S = (d, s)$ resident.

Generally, we can only calculate ρ by (i) numerically solving equation (1) until some time t_1 that should be sufficient to allow the resident population to relax to its attractor, (ii) numerically solving equation (3) or, equivalently, equation (2) in parallel with equation (1) starting from some positive initial condition at t_1, and (iii) estimating ρ from the average linear increase rate of $\ln(D^*)$ with time.

When the resident dynamics converge to a point equilibrium we can calculate ρ as the dominant eigenvalue of the corresponding constant matrix B. However, this is hard work and, in this case, there is an easier solution to the invasion problem as discussed in § 2 (d).

(d) *A stand-in for fitness in the case of population dynamical point equilibria*

In order to remove clutter, we shall adhere to the convention that 'individual' and 'disperser' refer to mutant individuals, dispersing mutants, etc.

In the model as formulated, individuals may live in very many different environments. They are either dispersing or they are in an (i, j) patch and, therefore, have i residents and $j - 1$ mutants impinging on them. This makes calculating the usual R_0, i.e. the mean lifetime number of children of a randomly chosen individual, a difficult task. The way out is to proceed not on an individual but on a colony basis, i.e. to work not from birth to births, but from dispersal event to dispersal events, on the rationale that in the usual structured metapopulation model all dispersers are equivalent.

We shall call the average number of dispersers produced by the (0 or 1) colonies founded by a newly born disperser R_m, m being from metapopulation. We calculate R_m in a number of steps. First, we observe that a newborn disperser has probability

$$\pi_i = \frac{\alpha^* s_{i,0}^* \widehat{p}_i}{\alpha^* \sum_{j=0}^{k-1} s_{j,0}^* \widehat{p}_j + \mu_D^*}, \tag{4}$$

of ending up in an $(i, 0)$ patch, turning it into an $(i, 1)$ patch. The newly founded colony then undergoes a Markovian stochastic population process until extinction. The states of this continuous-time Markov chain are the pairs (i, j), $j > 0$ and $i \geqslant 0$, $(i + j) \leqslant k$. We shall number these states according to the same scheme as used in § 2 (c). The corresponding probability vector $\boldsymbol{X}(a)$, where a is the age of the colony and $x_n(a)$ is the probability of finding the colony in state $n = L(i, j)$ at age a (with L the renumbering scheme defined in § 2 (c)), satisfies

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}a} = \widetilde{B}\boldsymbol{X}, \tag{5a}$$

$$\boldsymbol{X}(0) = \boldsymbol{Y}, \tag{5b}$$

with $y_n = \pi_i$ for $n = L(i, 1)$, $i = 0, \ldots, k - 1$ and $y_n = 0$ for all other n and with $\widetilde{B}$ constructed from B by removing the last row and column. (i, j) patches produce dispersers at a rate $j\lambda_{i,j}^* d_{i,j}^*$. We collect those rates in a row vector $\boldsymbol{A}$ in the usual manner, i.e. we set $j\lambda_{i,j}^* d_{i,j}^* = a_{L(i,j)}$ for $i = 0, \ldots, k - 1$, $j = 1, \ldots, k - i$. Then the general theory of continuous-time Markov chains tells us that

$$R_m = -\boldsymbol{A}\widetilde{B}^{-1}\boldsymbol{Y} \tag{6}$$

or, equivalently,

$$R_m = \boldsymbol{A}\boldsymbol{Z}, \tag{7}$$

with $\boldsymbol{Z}$ the solution of

$$\widetilde{B}\boldsymbol{Z} = -\boldsymbol{Y}. \tag{8}$$

Appendix B describes a robust numerical method for calculating R_m based on equations (7) and (8) for smallish (for example, $k < 50$) patch sizes. We derive an approximation for R_m applicable for larger patch sizes in § 3.

R_m is a function of two variables, the resident strategy S and the mutant strategy S^*, which we can express by writing $R_m(S, S^*)$. Consistency requires that $R_m(S, S) = 1$. This property was born out by numerical work on various special models, but we have not been able to prove it in general.

3. THE INFINITE PATCH SIZE CASE

(a) *Taking the limit for patch size going to infinity*

There exists a large body of theory for structured metapopulations in which the local population density x is treated as a continous variable (e.g. Gyllenberg *et al.* 1997). Biologically these models can be seen as useful limits for large patch sizes of models with discrete local population sizes. We proceed in this spirit. We shall argue heuristically in what manner our results simplify when we let the size of the patches (ω) move towards infinity. Moreover, we immediately concentrate on situations where the populations reside in a point equilibrium, as these are the only cases where we can arrive at analytical results. For fluctuating environments we have to go numeric, which comes close to using the model with a discrete structuring variable from which we started. (This statement is a bit facetious as there exist efficient numerical techniques for handling deterministic structured population models with continuous structuring variables in a more direct manner (De Roos & Metz 1991; De Roos *et al.* 1992).)

First, we consider the dynamics of the local resident densities $x = i/\omega$, i the number of residents and the correspondingly scaled disperser pool $M = D/\omega$ without mutants. In order to arrive at a continuum limit we have to make the biologically reasonable assumption that our individuals, be they mutants or residents, only react to local densities:

$$\lambda_i = \lambda(i/\omega), \quad \mu_i = \mu(i/\omega), \quad \gamma_i = \gamma(i/\omega),$$
$$d_i = d(i/\omega), \quad s_i = s(i/\omega). \tag{9}$$

After ω has gone to infinity and in between catastrophes, the resident dynamics satisfy

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = g(x, M) \triangleq [\lambda(x)(1 - d(x)) - \mu(x)]x + \alpha s(x)M, \qquad (10)$$

where τ is the time since the last catastrophe and

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \int_0^\infty \lambda(x)d(x)xp(x)\mathrm{d}x - \alpha \int_0^\infty s(x)p(x)\mathrm{d}x\ M - \mu_D M,$$
$$(11)$$

where p is the current probability density of the local population densities. This probability density can be calculated from a partial differential equation, as discussed by Metz & Diekmann (1986) and Gyllenberg & Hanski (1992) or using the integral equation approach discussed in Gyllenberg *et al.* (1997; also cf. Diekmann *et al.* 1998, 2001) using g from equation (10) as an ingredient.

Mutants start their career as single individuals. Therefore, if we consider how a mutant population takes off we have to deal with infinitesimally small values of M^*. In contrast to the resident case, there is no continuous stream of mutant dispersers into the patches, at least during the initial phase of the invasion process. We have to consider single arrival events. The resulting infinitesimally small local mutant population densities never grow to an appreciable size since local populations only have a finite time prior to being wiped out by a catastrophe. Therefore, even for infinite ω, we have to treat the local mutant population dynamics as a stochastic process. (The reason we did not have this discrepancy between mutant and resident dynamics in the case of finite ω is that the small patch size makes the local residents behave stochastically as well. In the case of finite ω it is only D^* which is very small relative to D. This means that the relative density of patches that have mutants in them is very small. However, when we concentrate on what happens in those patches, mutant and resident population sizes are of the same order of magnitude.)

The small size of the local mutant populations also has a helpful side. The presence of mutants leaves the local resident population unaffected. In the same vein, the population dynamical characteristics of the mutant individuals are unaffected by the local mutant density. Only the local value of x matters. Therefore, the local mutant numbers j develop according to a stochastic linear birth and death process with x-dependent, time-varying parameters.

(b) *Calculating R_m*

R_m can be calculated along similar lines as in the finite patch size case. Our closing example has been chosen so that it is possible to skip most of the calculations. However, we shall outline the full procedure as it is both conceptually relevant and a necessary ingredient in other applications.

First, we have to calculate the equilibrium values $\widehat{p}$ and $\widehat{M}$ for the probability density of the local resident population densities p and the disperser density M. Appendix C gives a recipe. We calculate the probability density π for the patches in which a settling mutant disperser will find itself from

$$\pi(x) = \frac{\alpha^* \widehat{p}(x) s^*(x)}{\alpha^* \int_0^\infty \widehat{p}(\xi) s^*(\xi) d\xi + \mu_D^*}. \qquad (12)$$

The resident density $y_x(a)$ surrounding a mutant colony for which the resident density at the moment of founding was x can be calculated from

$$\frac{\mathrm{d}y_x}{\mathrm{d}a} = g(y_x, \widehat{M}), \qquad (13a)$$

$$y_x(0) = x. \qquad (13b)$$

As a consequence of the independence of the individuals in the mutant colony the average size m_x of that colony satisfies the differential equation

$$\frac{\mathrm{d}m_x}{\mathrm{d}a} = [\lambda^*(y_x)(1 - d^*(y_x)) - \mu^*(y_x) - \gamma(y_x)]m_x, \qquad (14a)$$

$$m_x(0) = 1, \qquad (14b)$$

where we now include the possibility that the colony has been eradicated by a catastrophe. The expected reproductive output at age a from such a colony equals $\lambda^*(y_x(a))d^*(y_x(a))m_x(a)$. Therefore, the expected number of dispersers produced by a colony founded by a mutant disperser entering at resident density x equals

$$R(x) = \int_0^\infty \lambda^*(y_x(a))d^*(y_x(a))m_x(a)\mathrm{d}a, \qquad (15)$$

and

$$R_\mathrm{m} = \int_0^\infty R(x)\pi(x)\mathrm{d}x. \qquad (16)$$

Appendix D describes a simple procedure for calculating R_m using existing packages for solving differential equations.

4. EXAMPLE: CALCULATING EVOLUTIONARILY STABLE DISPERSAL STRATEGIES

We have already included two explicit strategy vectors in our basic model formulation, $(d_i)_{i=1, \ldots, k-1}$, the probabilities that a newborn disperses when born in a patch with population size i and $(s_j)_{j=0, \ldots, k-1}$, the probabilities that a disperser stays on encountering a patch with population size j. Here we shall only consider the simpler infinite patch size case. In that case the strategy consists of two functions d and s of the continuous variable x. In order to simplify calculating the ESS, which is denoted as $(\widehat{d}, \widehat{s})$, we introduce some biologically reasonable assumptions on the other model ingredients. We assume that both the birth rate λ and $\lambda - \mu - \gamma$ are non-increasing continuous functions of x, $\lambda(0) > \mu(0) + \gamma(0)$ and there exists a unique positive $\widetilde{x}$ such that

$$\lambda(\widetilde{x}) = \mu(\widetilde{x}) + \gamma(\widetilde{x}). \qquad (17)$$

We obtain the ESS by maximizing $R_\mathrm{m}[(d,s), (d^*, s^*)]$ in (d^*, s^*) and then setting mutant equal to resident (so that $R_\mathrm{m} = 1$).

Due to the lack of memory of the disperser state a just-arriving immigrant in a patch with resident density x has a future indistinguishable from a newly born individual at the same value of x. Therefore, $\widehat{d}(x) = 1 - \widehat{s}(x)$ for all x where $(\widehat{d}(x), \widehat{s}(x))$ is the unique maximizer of R_m in the $(d^*(x), s^*(x))$ direction. It will turn out that, at such x,

either $\widehat{d}(x) = 1 - \widehat{s}(x) = 1$ or $\widehat{d}(x) = 1 - \widehat{s}(x) = 0$. Where uniqueness fails there is a one-dimensional continuum of maximizing values; below we shall see that this only happens at a single special value of x.

Since mutant individuals reproduce and die independently, for a newborn stayer the expected number of dispersing descendants produced by itself and the clan of all its within-patch descendants also equals $R(x)$. At population dynamical equilibrium a disperser on average produces, through starting a colony, $R_{\mathrm{m}} = 1$ new disperser. A mutant who, at some x, behaves in a manner which lets it and its clan produce, on average, more future dispersers than the resident and everywhere else do not do worse will invade. Thus, a strategy that has stayers at local densities where staying results in a net loss ($R(x) < 1$) can always be invaded and, therefore, cannot be an ESS. An analogous argument applies to a strategy that has leavers at local densities where staying would result in a net gain ($R(x) > 1$). (This is the usual marginal value type of argument, but is based on a fitness measure geared to a metapopulation situation as well as taking account of density dependence.)

$R(x)$ can be calculated from equations (14) and (15). In Appendix E we prove that, independent of the details of the dispersal strategy, $R(x) < 1$ in any patches with $x > \widetilde{x}$. Therefore, in the ESS every newborn should disperse whenever $x > \widetilde{x}$. We also prove in Appendix E that, if we consider only strategies for which every newborn disperses whenever $x > \widetilde{x}$, then $R(x) > 1$ in patches with $x < \widetilde{x}$. Therefore, in the ESS newborns should stay as long as $x < \widetilde{x}$.

As a consequence, at the ESS $0 < \widehat{p}(x) < \infty$ for $0 \leqslant x < \widetilde{x}$ and $\widehat{p}(x) = 0$ for $\widetilde{x} < x$. Moreover, $\widetilde{x}$ is reachable from 0 in finite time. Therefore, $\widehat{p}$ contains a delta function component at $\widetilde{x}$ representing a concentrated probability mass.

Consistency requires that the local population growth rate of the resident precisely at $\widetilde{x}$ is

$$g(\widetilde{x}, \widehat{M}) = [\lambda(\widetilde{x})(1 - \widehat{d}(\widetilde{x})) - \mu(\widetilde{x})]\widetilde{x} + \alpha\widehat{s}(\widetilde{x})\widehat{M} = 0, \quad (18)$$

for, if $g(\widetilde{x}, \widehat{M})$ were larger than 0 at the ESS, the local population density would grow beyond $\widetilde{x}$, which is incompatible with the fact that, beyond $\widetilde{x}$, all newborns leave, with an analogous argument applying on the other side of $\widetilde{x}$.

A biologically more realistic phrasing of the previous arguments runs as follows. Below $\widetilde{x}$ individuals should stay and above $\widetilde{x}$ they should leave. At $\widetilde{x}$ it is unclear what they should do. In practice, they will perceive the surrounding population density with some small error. Thus, some will leave a little too early or a little too late. If by chance many were to err on the late side the local population density would increase further beyond $\widetilde{x}$ forcing them to leave anyway. This natural feedback loop would (i) produce a very steep hump in $\widehat{p}$ around $\widetilde{x}$, with $\widehat{p} = 0$ somewhat further beyond $\widetilde{x}$, and (ii) produce a pattern of leaving over the small x interval under the hump, which when looked at through foggy glasses would be indistinguishable from equation (18).

At the ESS $R(\widetilde{x}) = 1$ independent of the value of $\widehat{d}(\widetilde{x})$. Apart from equation (18) there is no further constraint on $\widehat{d}(\widetilde{x})$ and $\widehat{s}(\widetilde{x})$.

The final conclusion is that, overall, the ESS has the pattern of a bang-bang control: $\widehat{s}(x) = 1$ and $\widehat{d}(x) = 0$ for all $x < \widetilde{x}$ and $\widehat{s}(x) = 0$ and $\widehat{d}(x) = 1$ for all $x > \widetilde{x}$. At $x = \widetilde{x}$ the decisions are no longer deterministic. At this value of x the ESS is also non-unique: there is a one-dimensional continuum of ESSs consisting of all pairs $(\widehat{s}(\widetilde{x}), \widehat{d}(\widetilde{x}))$, $0 \leqslant \widehat{s}(\widetilde{x}), \widehat{d}(\widetilde{x}) \leqslant 1$, satisfying equation (18). At the ESS the population density in a sizeable fraction of patches is the same, namely $\widetilde{x}$; the densities in the remaining patches vary between 0 and $\widetilde{x}$.

The prediction about the equilibrium distribution of the population densities suggests an easy field test. Unfortunately, this prediction hinges on the simplifying assumption of equal patch qualities. If patch qualities differ, the ES dispersal strategy stays of the bang-bang kind, but the local value of $\widetilde{x}$ is dependent on the patch quality, resulting in an ideal free arrangement for the 'filled' patches. Within a set of patches of the same quality, again a sizeable fraction of patches should be filled, supporting the population density $\widetilde{x}$ that corresponds to the local quality, while the densities in the remaining patches in the set should vary between 0 and that $\widetilde{x}$. If no independent estimate of patch quality is available, the theory can be tested by looking at whether, after local catastrophes, the local dispersal rate drops back to zero, eventually to bounce back, each time at approximately the same value of x and to approximately its former non-zero value, which is such that, from then on, it keeps the local population density approximately constant, that is until the occurrence of the next catastrophe.

5. DISCUSSION

We gave a fitness definition for metapopulations consisting of a large number of equally coupled patches. The trick in deriving an appropriate fitness concept in a metapopulation context is to replace individuals by local colonies together with their local environment as the unit of calculation. This trick is only technical; we are dealing with a strictly individual-based fitness concept all along. The colony viewpoint is only introduced in order to keep track of the varying environments in which individuals find themselves. (This way we also account for kin selection effects.) In addition, we devised a quantity that stands in the same mathematical relation to metapopulation fitness as R_0 stands to ordinary fitness: R_{m} is only defined for population dynamical point equilibria and then the sign of $\ln(R_{\mathrm{m}})$ equals the sign of the metapopulation fitness. Basically, R_{m} describes the fate of local colonies by following their life cycle from a newborn disperser to the dispersers produced by the colony that it may found. Algorithms for calculating R_{m} for smallish patch sizes are described in Appendices A and B. In addition, we give an approximation of R_{m} for largish patch sizes in terms of some formidable looking integrals. Appendices C and D describe how the integrals can be calculated using readily available packages for the solution of differential equations.

In §4 we showed how the newly developed concepts can be used to solve a problem posed by Olivieri & Gouyon (1998): What is the conditional dispersal strategy when individuals react to the local population density? This strategy can be calculated through an adapted

marginal value argument, phrased in terms of the newly developed fitness measure and taking account of the inherent density dependence, which necessarily makes the R_m of any resident type equal to one. The result is akin to an ideal free distribution: the ES dispersal behaviour makes the patches reside at densities where the expected contribution of the individuals to future generations are exactly equal as much as possible, independent of whether their offspring stay at home or disperse. Only patches that have recently been subject to a catastrophic extinction have lower population densities so that the individuals in them have higher expected contributions to future generations. However, in order to reach such patches dispersers have to pay the penalty of potentially dying during the process, which equalizes the expected fitness gain precisely.

At least three earlier papers have dealt with density-dependent dispersal. All three considered discrete-time models, with dispersal in a single pulse after reproduction. Ezoe & Iwasa (1997) assumed that $k = 1$, as opposed to $k = \infty$ in our model. Dispersal reduces the number of remaining offspring in a deterministic manner, which is only possible if there is a dependence between the dispersal decisions of the offspring, for example, since the decision is taken by the mother or when the number of offspring is very large. Reproduction fluctuates randomly in space and time. The ES dispersal fraction is zero when the number of offspring is below a threshold and otherwise sets this number back to the threshold density. Jánosi & Scheuring (1997) used a deterministic coupled map lattice for their population model. The density fluctuations necessary for the profitability of dispersal come from a non-point attractor. They restricted their attention to strategies where individuals do not disperse below a threshold density and otherwise disperse in a complicated, dependent manner which reduces the population density to somewhat below the threshold density. Their fig. 5 shows a unique evolutionarily attracting ES threshold, which, however, was not robust: in principle there exist slight changes in model specification that transform the ESS into a branching point (Metz *et al.* 1996; Geritz *et al.* 1998). The model of Travis *et al.* (1999) is closest to ours in spirit except that it does not incorporate any environmental stochasticity and that the strategies are restricted to functions that are zero below a threshold density and increase linearly thereafter. The restriction to a small subset of possible dispersal rules makes the numerical results of Jánosi & Scheuring (1997) and of Travis *et al.* (1999) incommensurable with ours. The analytical result of Ezoe & Iwasa (1997) conforms with our result that, whenever possible, the ES dispersal rule fixes the local population density at a single threshold value.

The particular example we used has the advantage that the ESS is monomorphic. The arguments used for its calculation also show that it has a fair domain of evolutionary attraction. The techniques of adaptive dynamics (e.g. Geritz *et al.* 1998) together with the procedures for calculating R_m also make it possible to determine whether a sequence of quasi-monomorphic substitutions eventually leads to evolutionary branching. After such a branching event mutant fitness can still be defined in a similar manner and so can R_m, but the state space of the local

mutant clans has to be expressed in terms of two resident population sizes i_1 and i_2 in addition to the mutant population size j (or more population sizes if higher degrees of polymorphism occur). In the finite patch size case this does away with our trick in Appendix A for calculating the population dynamical equilibrium. Moreover, the numerics in Appendix B for calculating R_m quickly gets out of hand. We eagerly await further practical tricks for tackling such cases. At least we now have an understanding of the necessary fundamentals.

APPENDIX A. HOW TO CALCULATE THE $\widehat{p}_i$

In order to calculate $\widehat{p}_i$ we start by solving $F(\widehat{D}) = 0$, e.g. using a bisection method, where F is defined by the following algorithm (if $F(D) < 0$ for all $D > 0$ the metapopulation is not viable).

(i) Supply a value of D.

(ii) The next step is to calculate two sequences of numbers to be called $\widetilde{p}_{1,i}$ and $\widetilde{p}_{2,i}$, $i = 0, \ldots, k$. The calculation is started by setting $\widetilde{p}_{1,0} = 0$ and $\widetilde{p}_{1,1} = 1$ and $\widetilde{p}_{2,0} = 1$ and $\widetilde{p}_{2,1} = 0$.

(iii) Successively calculate the following $\widetilde{p}_{1,i}$, $i = 2, \ldots, k$, using
$$\widetilde{p}_{1,i+1} = \{[i\mu_i + \alpha D s_i + i\lambda_i(1 - d_i) + \gamma_i]\widetilde{p}_{1,i} - [\alpha D s_{i-1} + (i-1)\lambda_{i-1}(1 - d_{i-1})]\widetilde{p}_{1,i-1}\}/\{(i+1)\mu_{i+1}\}.$$

(iv) Calculate the $\widetilde{p}_{2,i}$ in exactly the same manner as the $\widetilde{p}_{1,i}$.

(v) Calculate $P_1 = \sum_{i=0}^{k} \widetilde{p}_{1,i}$ and $P_2 = \sum_{i=0}^{k} \widetilde{p}_{2,i}$.

(vi) Calculate $Q_1 = \sum_{i=1}^{k} \gamma_i \widetilde{p}_{1,i}$ and $Q_2 = \sum_{i=1}^{k} \gamma_i \widetilde{p}_{2,i}$.

(vii) Calculate $W = (\mu_1 + Q_1)P_2 + (\alpha D s_0 - Q_2)P_1$.

(viii) Calculate $u_1 = (\alpha D s_0 - Q_2)/W$ and $u_2 = (\mu_1 + Q_1)/W$.

(ix) Calculate the numbers $\widehat{p}_i = u_1 \widetilde{p}_{1,i} + u_2 \widetilde{p}_{2,i}$.

(x) Calculate $F(D)$ as
$$F(D) = \sum_{i=1}^{k} i\lambda_i d_i \widehat{p}_i - \alpha \sum_{i=0}^{k-1} \widehat{p}_i s_i D - \mu_D D.$$

From the equilibrium $\widehat{D}$ of D, we can calculate the equilibrium values $\widehat{p}_i$ of the p_i by the same rules as before with $\widehat{D}$ substituted for D.

APPENDIX B. HOW TO CALCULATE R_m WHEN ω IS FINITE

The renumbering scheme L used in §2(c,d) had bringing out the classical vector–matrix structure of the procedures as its only purpose. In order to calculate $\mathbf{Z}$ from equation (8) it is easier to return to the original numbering scheme in terms of i and j (the numbers of residents and mutants in the patch, respectively). We do this by setting $u_{i,j} = z_{L(i,j)}$. The $u_{i,j}$ can be interpreted as the mean time the local mutant colony stays in state (i,j) before extinction. The components of the matrix $\widetilde{B}$ we denote as

$$c_{i,j}^a = (j+1)\mu_{i,j+1}^*,$$

$$c_{i,j}^l = \alpha\widehat{D}s_{i-1,j} + (i-1)\lambda_{i-1,j}(1 - d_{i-1,j}),$$

$$c_{i,j}^c = -[i\mu_{i,j} + j\mu_{i,j}^* + \alpha\widehat{D}s_{i,j} + i\lambda_{i,j}(1 - d_{i,j})$$
$$\qquad + j\lambda_{i,j}^*(1 - d_{i,j}^*) + \gamma_{i,j}^*],$$

$$c_{i,j}^r = (i+1)\mu_{i+1,j},$$

$$c_{i,j}^b = (j-1)\lambda_{i,j-1}^*(1 - d_{i,j-1}^*)$$

for $j > 1$. (The indices a, l, c, r, and b stand for above, left, centre, right and below, respectively. Now let

$$v_{i,j} = -c_{i,j}^c u_{i,j},$$

so that

$$R_m = \sum_{i=0}^{k-1} \sum_{j=1}^{k-i} \frac{j\lambda_{i,j}^* d_{i,j}^*}{-c_{i,j}^c} v_{i,j}.$$

The $v_{i,j}$ can be interpreted as the number of times the colony passes through state (i,j). They can be calculated as the limit $v_{i,j}(\infty)$ from the recurrence

$$v_{i,j}(n+1) = \frac{c_{i,j}^a}{-c_{i,j+1}^c} v_{i,j+1}(n) + \frac{c_{i,j}^l}{-c_{i-1,j}^c} v_{i-1,j}(n)$$
$$\qquad + \frac{c_{i,j}^r}{-c_{i+1,j}^c} v_{i+1,j}(n) + \frac{c_{i,j}^b}{-c_{i,j-1}^c} v_{i,j-1}(n) \qquad \text{(B1)}$$

for $i \geqslant 0, j \geqslant 1$ and $i + j \leqslant k$, together with the convention

$$c_{i,1}^b = 1,$$

$$c_{i+1,0}^c = -1,$$

and the boundary conditions $v_{-1,j} = 0$, $v_{i,0} = \pi_i$ and $v_{i,j} = 0$ for $i + j = k + 1$, starting from any possible initial condition. If one has no better initial guess one can take $v_{i,j}(0) = 0$. However, usually the calculation will be part of an ESS calculation or of a scheme where the parameter dependence of that ESS is analysed. In that case one generally already has a fair initial guess for the $v_{i,j}$ from some earlier stage of the calculation which can be improved using equation (B1).

APPENDIX C. HOW TO CALCULATE $\widehat{p}(x)$

(a) *The general procedure*

In order to calculate $\widehat{p}(x)$ we first calculate the function $\widetilde{p}$ of the two variables x and M as

$$\widetilde{p}(x, M) = \frac{C}{g(x, M)} \exp\left[-\int_0^x \frac{\gamma(\xi)}{g(\xi, M)} d\xi\right] \quad \text{for } 0 \leqslant x < \widehat{x}(M)$$
$$= 0 \qquad\qquad\qquad\qquad \text{for } \widehat{x}(M) < x,$$

with

$$C = \left(\int_0^\infty \frac{1}{g(x, M)} \exp\left[-\int_0^x \frac{\gamma(\xi)}{g(\xi, M)} d\xi\right] dx\right)^{-1},$$

where $\widehat{x}(M)$ is the positive solution of

$$g(\widehat{x}, M) = 0,$$

provided such a solution exists (in which case it is unique), otherwise $\widehat{x}(M) = \infty$.

The next step is to calculate $\widehat{M}$ by (numerically) solving the equation

$$\int_0^\infty \lambda(x)d(x)x\widetilde{p}(x,\widehat{M})dx - \alpha\int_0^\infty s(x)\widetilde{p}(x,\widehat{M})dx\widehat{M}$$
$$\qquad - \mu_D\widehat{M} \triangleq G(\widehat{M}) = 0.$$

(If $G(M) < 0$ for all $M > 0$ the metapopulation is not viable.) Finally, we calculate $\widehat{p}$ as

$$\widehat{p}(x) = \widetilde{p}(x, \widehat{M}).$$

In practice it will usually be necessary to evaluate the various integrals numerically.

(b) *An efficient recipe for calculating the integrals*

For a given fixed M there exists a monotone relationship (to be called X) between the age of the local resident population τ, which is defined as the time since a patch underwent its last catastrophe and the present resident population density x. The existence of such a relationship makes it possible to calculate the integrals using a standard package for solving differential equations. The equations to be solved are (with the two arguments τ and M suppressed; the X, Y, Z, V and W below bear no relation to the same symbols from the main text)

$$\frac{dX}{d\tau} = g(X, M) \qquad X(0) = 0,$$

$$\frac{dY}{d\tau} = -\gamma(X)Y, \qquad Y(0) = 1,$$

$$\frac{dZ}{d\tau} = Y, \qquad Z(0) = 0,$$

$$\frac{dV}{d\tau} = \lambda(X)d(X)XY, \qquad V(0) = 0,$$

$$\frac{dW}{d\tau} = s(X)Y, \qquad W(0) = 0.$$

X describes the development of the local population size and Y the survival of the colony. The other three quantities are only calculated in order to find their values for large τ, which are needed as a normalization factor and in the equation for $\widehat{M}$. Given the solution of these differential equations G can be calculated as

$$G(M) = V(\infty)/Z(\infty) - [\alpha W(\infty)/Z(\infty) + \mu_D]M.$$

It takes a long integration to arrive at $\tau = \infty$. However, in practice one can replace the tail of the integration by

an analytical approximation: as time proceeds $X(\tau)$ converges to $\widehat{x}(M)$ which is defined as the solution of $g(\widehat{x}, M) = 0$. When $X(\tau)$ has come sufficiently close to $\widehat{x}(M)$, say at $\tau = T$, one can stop integrating and write

$$V(\infty) = V(T) + \lambda(\overline{X})d(\overline{X})\overline{X}\,\Upsilon(T)/\gamma(\overline{X}),$$

$$W(\infty) = W(T) + s(\overline{X})\,\Upsilon(T)/\gamma(\overline{X})$$

and

$$\mathcal{Z}(\infty) = \mathcal{Z}(T) + \Upsilon(T)/\gamma(\overline{X}),$$

where $\overline{X}$ is either $X(T)$ or $\widehat{x}(M)$. The discrepancy between the two cases can be used to judge whether one has chosen T sufficiently large.

The equilibrium colony age distribution $\widehat{q}$ corresponds to

$$\widehat{q}(\tau) = \widehat{\Upsilon}(\tau)/\widehat{\mathcal{Z}}(\infty),$$

where a 'hat' means that the quantity has been evaluated at $M = \widehat{M}$. The stationary colony size distribution is calculated by transforming from population age to size:

$$\widehat{p}(\widehat{X}(\tau)) = \frac{\widehat{q}(\tau)}{g(\widehat{X}(\tau), \widehat{M})}.$$

APPENDIX D. HOW TO CALCULATE R_{m} WHEN ω IS INFINITE

In order to calculate R_{m} in practice, it is again easier to revert to a representation in terms of local population ages. We start calculating a quantity U from

$$\frac{\mathrm{d}U}{\mathrm{d}\tau} = f^*(X)U, \qquad U(0) = 1,$$

with

$$f^*(X) = [\lambda^*(\widehat{X})(1 - d^*(\widehat{X})) - \mu^*(\widehat{X}) - \gamma(\widehat{X})],$$

and $\widehat{X}(\tau)$ the quantities already calculated in Appendix C § (b). Given U we can calculate m from equation (14b) as $m_{\widehat{X}(\tau)}(a) = U(\tau + a)/U(\tau)$.

Let $\widehat{\widetilde{x}}$ be the solution of $g(\widehat{\widetilde{x}}, \widehat{M}) = 0$. If $f^*(\widehat{\widetilde{x}}) > 0$ then $R_{\mathrm{m}} = \infty$ and the mutant can invade. Under some reasonable monotonicity and smoothness assumptions the same holds good when $f^*(\widehat{\widetilde{x}}) = 0$. So we can concentrate on the case where $f^*(\widehat{\widetilde{x}}) < 0$.

In order to calculate $R(x)$ we introduce a quantity Q to be calculated from

$$\frac{\mathrm{d}Q}{\mathrm{d}\tau} = \lambda^*(\widehat{X})d^*(\widehat{X})U, \qquad Q(0) = 0.$$

This allows us to calculate $R(\widehat{X}(\tau)) = [Q(\infty) - Q(\tau)]/U(\tau)$. In practice, for large τ we can again replace Q with the analytical approximation

$$Q(\infty) \approx Q(T) - \lambda^*(\overline{X})d^*(\overline{X})U(T)/f^*(\overline{X})$$

and

$$Q(\tau) \approx Q(\infty) - [Q(\infty) - Q(T)]\exp[f^*(\overline{X})(\tau - T)],$$

where T and $\overline{X}$ are defined as in Appendix C § (b), but now with $\widehat{M}$ substituted for M.

It remains to integrate $R(\widehat{X}(\tau))$ with the patch age distribution calculated in Appendix C. To this end, we introduce three helper quantities A, B and C to be calculated from

$$\frac{\mathrm{d}A}{\mathrm{d}\tau} = \frac{s^*(\widehat{X})\widehat{\Upsilon}}{U}, \qquad A(0) = 0,$$

$$\frac{\mathrm{d}B}{\mathrm{d}\tau} = \frac{Qs^*(\widehat{X})\widehat{\Upsilon}}{U}, \qquad B(0) = 0,$$

$$\frac{\mathrm{d}C}{\mathrm{d}\tau} = s^*(\widehat{X})\widehat{\Upsilon}, \qquad C(0) = 0,$$

with $\widehat{\Upsilon}(\tau)$ the quantities calculated in Appendix C § (b). In principle this allows us to calculate R_{m} as

$$R_{\mathrm{m}} = \alpha^* \frac{\lim_{\tau \to \infty}[Q(\infty)A(\tau) - B(\tau)]}{\alpha^*C(\infty) + \mu_D^*\widehat{\mathcal{Z}}(\infty)},$$

with $\widehat{\mathcal{Z}}(\infty)$ the quantity calculated in Appendix C § (b). The reason for not writing the numerator as $Q(\infty)A(\infty) - B(\infty)$ is that both $A(\infty)$ and $B(\infty)$ are infinite. The numerical trouble that this generates is avoided by using the approximation

$$\lim_{\tau \to \infty}[Q(\infty)A(\tau) - B(\tau)] \approx Q(\infty)A(T) - B(T)$$
$$+ \frac{Q(\infty) - Q(T)}{U(T)}\frac{s^*(\overline{X})\widehat{\Upsilon}(T)}{\gamma(\overline{X})}.$$

APPENDIX E. SOME HELPFUL INEQUALITIES

We start from a more explicit expression for $R(x)$ calculated from equations (14) and (15):

$$R(x) = \int_0^\infty \lambda(y_x(a))d^*(y_x(a)) \times \exp\left[\int_0^a [\lambda(y_x(\tau))(1 - d^*(y_x(\tau)))\right.$$
$$\left. - \mu(y_x(\tau)) - \gamma(y_x(\tau))]\mathrm{d}\tau\right]\mathrm{d}a, \qquad (\mathrm{E1})$$

where y_x should still be calculated from equation (13).

The general theory of structured population models tells us that, for our model, $\widehat{p}(x) > 0$ if and only if $g(x, \widehat{M}) \geqslant 0$. There is no way in which a local population can ever reach values of x for which $g < 0$. Therefore, we can safely assume that in equation (E1) $y_x \geqslant x$. Therefore, $R(x) \leqslant H(x, x)$ always with H defined by

$$H(z, x) \triangleq \int_0^\infty \lambda(z)d^*(y_x(a)) \exp\left[\int_0^a [\lambda(z)(1 - d^*(y_x(\tau)))\right.$$
$$\left. - \mu(z) - \gamma(z)]\mathrm{d}\tau\right]\mathrm{d}a.$$

Together with our assumptions about λ, μ and γ, equation (17) guarantees that for $x > \widetilde{x}$

$$R(x) \leqslant H(x, x) < H(\widetilde{x}, x) = 1.$$

Therefore, in the ESS every newborn disperses whenever $x > \widetilde{x}$. As a consequence $\widehat{p}(x) = 0$ for all $x > \widetilde{x}$.

Now consider what happens at values of $x \leqslant \widetilde{x}$. Using the fact that x never grows beyond $\widetilde{x}$ we can conclude that, at the ESS, for $x < \widetilde{x}$

$$R(x) > H(\widetilde{x}, x) = 1.$$

REFERENCES

De Roos, A. M. & Metz, J. A. J. 1991 Towards a numerical analysis of the escalator boxcar train. In *Differential equations with applications in biology, physics and engineering* (ed. J. A. Goldstein, F. Kappel & W. Schappacher), pp. 91–113. New York: Marcel Dekker.

De Roos, A. M., Diekmann, O. & Metz, J. A. J. 1992 Studying the dynamics of structured population models: a versatile technique and its application to *Daphnia. Am. Nat.* **139**, 123–147.

Dieckmann, U. & Law, R. 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579–612.

Dieckmann, U., Law, R. & Metz, J. A. J. 2000 *The geometry of ecological interactions: simplifying spatial complexity.* Cambridge University Press.

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.

Diekmann, O., Gyllenberg, M., Metz, J. A. J. & Thieme, H. R. 1994 The 'cumulative' formulation of (physiologically) structures population models. In *Evolution equations, control theory, and biomathematics* (ed. P. Clément & G. Lumer), pp. 145–154. New York: Marcel Dekker.

Diekmann, O., Gyllenberg, M., Metz, J. A. J. & Thieme, H. R. 1998 On the formulation and analysis of general deterministic structured population models. I. Linear theory. *J. Math. Biol.* **36**, 349–388.

Diekmann, O., Gyllenberg, M., Huang, H., Kirkilionis, M., Metz, J. A. J. & Thieme, H. R. 2001 On the formulation and analysis of general deterministic structured population models. II. Nonlinear theory. *J. Math. Biol.* (In the press.)

Doebeli, M. & Ruxton, G. D. 1997 Evolution of dispersal rates in metapopulation models: branching and cyclic dynamics in phenotype space. *Evolution* **51**, 1730–1741.

Ezoe, H. & Iwasa, Y. 1997 Evolution of condition dependent dispersal: a genetic algorithm search for the ESS reaction norm. *Res. Popul. Ecol.* **39**, 127–137.

Geritz, S. A. H., Metz, J. A. J., Kisdi, É. & Meszéna, G. 1997 Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **78**, 2024–2027.

Geritz, S. A. H., Kisdi, É., Meszéna, G. & Metz, J. A. J. 1998 Evolutionarily singular strategies and the adaptive growth and granching of the evolutionary tree. *Evol. Ecol.* **12**, 35–57.

Gyllenberg, M. & Hanski, I. 1992 Single-species metapopulation dynamics: a structured model. *Theor. Pop. Biol.* **42**, 35–62.

Gyllenberg, M. & Metz, J. A. J. 2001 On fitness in structured metapopulations. *J. Math. Biol.* (In the press.)

Gyllenberg, M., Hanski, I. & Hastings, A. 1997 Structured metapopulation models. In *Metapopulation dynamics: ecology, genetics and evolution* (ed. I. Hanski & M. Gilpin), pp. 93–122. London: Academic Press.

Hanski, I. 1999 *Metapopulation ecology.* Oxford University Press.

Hanski, I. & Gilpin, M. 1997 *Metapopulation dynamics: ecology, genetics and evolution.* London: Academic Press.

Hanski, I. & Gyllenberg, M. 1993 Two general metapopulation models and the core-satellite species hypothesis. *Am. Nat.* **142**, 17–41.

Hastings, A. 1983 Can spatial variation alone lead to selection for dispersal? *Theor. Pop. Biol.* **24**, 244–251.

Heesterbeek, J. A. P. 1992 R_0. PhD thesis, Leiden University, The Netherlands.

Holt, R. D. & McPeek, M. A. 1996 Chaotic population dynamics favors the evolution of dispersal. *Am. Nat.* **148**, 709–718.

Jánosi, M. J. & Scheuring, I. 1997 On the evolution of dispersal in a spatially structured population model. *J. Theor. Biol.* **187**, 397–408.

Metz, J. A. J. & Diekmann, O. (ed.) 1986 *The dynamics of physiologically structured populations.* Berlin: Springer.

Metz, J. A. J., Nisbet, R. M. & Geritz, S. A. H. 1992 How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol.* **7**, 198–202.

Metz, J. A. J., Geritz, S. A. H., Meszéna, G., Jacobs, F. J. A. & Van Heerwaarden, J. S. 1996 Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and spatial structures of dynamical systems* (ed. S. J. van Strien & S. M. Verduyn Lunel), pp. 183–231. Amsterdam: North-Holland.

Olivieri, I. & Gouyon, P.-H. 1998 Evolution of migration rate and other traits: the metapopulation effect. In *Metapopulation dynamics: ecology, genetics and evolution* (ed. I. Hanski & M. Gilpin), pp. 293–323. London: Academic Press.

Parvinen, K. 1999 Evolution of migration in a metapopulation. *Bull. Math. Biol.* **61**, 531–550.

Ronce, O., Perret, F. & Olivieri, I. 2000 Evolutionarily stable dispersal rates do not always increase with local extinction rates. *Am. Nat.* **155**, 485–496.

Ruxton, G. D. & Rohani, P. 1999 Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony. *J. Anim. Ecol.* **67**, 530–539.

Travis, J. M. J., Murrell, D. & Dytham, C. 1999 The evolution of density-dependent dispersal *Proc. R. Soc. Lond* B **266**, 1837–1842.